An Exploration based on Multifarious Video Copy Detection Strategies

Suman Elizabeth Daniel[1], Binu A.[2]
[1] Post Graduate Student Dept. of IT in Rajagiri School of Engineering and Technology, Kochi, India
Email: sumanedan@gmail.com
[2] Asst. Professor Dept. of IT in Rajagiri School of Engineering and Technology, Kochi, India
Email: binu_a@rajagiritech.ac.in

***Abstract* —We co-exist in an era, where tonnes and tonnes of videos are uploaded every day. Video copy detection has become the need for the hour as most of them are user generated Internet videos through popular sites such as YouTube. It acts as a medium to restrain piracy and prove whether the contents are legitimate. The usual procedure adopted in video copy detection techniques include discovering whether a query video is copied from a database of videos or not. This paper acquaints different Video copy detection techniques that have been adopted to ensure robust and secure videos along some applications of video fingerprinting.**

***Index Terms* — Video copy Detection, Fingerprints, Video Fingerprinting, Watermarking, Content based Video copy Detection, TIRI-DCT algorithm**

I. Introduction

In the present day, video copy detection has been focused extensively due to the huge number of videos that are being expelled in the Internet. Most of these videos are identified as manipulated versions of the existing videos or illegitimate copies which are being largely distributed in the Internet. The need for adoption of copy detection techniques arises for the identification of legal video content along with their management [8]. The complexity of video as digital media made advancements in this behalf slower. Moreover since videos are available in different formats, for efficient copy detection it was realized that copy detection should be based on the content of the video rather than any other means. Thus the concept of fingerprinting or Content-Based video copy detection gained large scale popularity.

II. Overview of Fingerprints

A video fingerprint is a signature [7] of the video which is unique and can be applied for curbing illegal violations of the video. This signature is derived from the content of the video [4]. A video fingerprint should possess following remarkable properties which would make it useful for adoption in copy detection process for a diverse set of videos:

A. Robust Nature

A manipulated video would have a similar fingerprint to that of the legitimate video.

B. Independent Pair Wise

Two dissimilar video would have different fingerprint.

C. Efficient Database Search

The Fingerprint can be applied efficiently for searching in large database of media [15].

D. Lower Complexity

The fingerprint generation algorithm should be accomplished in lower computational time so that video fingerprints could be computed quickly.

E. Compact

The fingerprint generated should be in miniature size as compared to content of the video making it comfortable to store in a media database.

III. Classification of Fingerprints

The Fingerprints can be classified into four major groups which are:

A. Color-Space Based Fingerprints

These depend on color or gray level properties of frames such as hue, saturation etc. and are obtained as histograms of colors in particular regions over specific time/space within a video. But this cannot be integrated with black and white videos [3]

B. Temporal Fingerprints

Here the key concept lies on the differences among the frames or the order of frames. They are derived from video sequence characteristics over time. They seem to have high performance for long video clips where as short video clips could not project much temporal information making it inapplicable for short videos.

C. Spatial Fingerprints

They are the features extracted from each and every frame or from a key frame of a video. They are being widely used for video and image fingerprinting. Here each pixel of a video frame is considered separately based on their location. Spatial fingerprint can be subdivided to global and local fingerprint where global fingerprints represent global properties of a frame or a subsection of it. Local fingerprint concentrates on local information of points within a frame.

D. Spatial-Temporal Fingerprints

Here features are derived based on local variation in both space and time. It is focused on differential of luminance of

DOI: 03.LSCS.2013.5. 47

grid partitions in spatial and temporal regions [10].

IV. Existing Video Copy Detection Techniques

The existing techniques for video copy detection can be categorized as:

A. Watermarking approach.
B. Image based approach for video copy detection.
C. Content based copy detection system.

A. Watermarking Approach

In this method a video stream in incorporated with information or watermarks which paves way detection of video copies. The watermarks may be visible [text or logo of producer or broadcaster] or may be invisible which cannot be perceived by human eye. Watermarking uses 3D-DCT algorithm whose binarization phase act as its major drawback [14]. A common threshold is needed for binarization phase which is far from optimal as different frames will have different frequencies. Watermarks are more prone to visual transformations such as re-encoding, change of resolution/ bit rate etc. Two major drawbacks that came forward are ensuring legitimate content difficulty and illegal attacks. The difficulty in ensuring legitimate content is because watermarks are integrated to original video before copies are made. So they cannot be applied to already circulated videos and thus watermarking approach could not pose an overall solution in this regard. There could also be illegal attacks on the video such that the watermark on the particular frame is compromised. In such a case there exists no alternative approach for video copy detection. [5]

B. Image Based Approach for Video Copy Detection

The Image based approach is centered on key frame based video copy detection in which individual frames are diagnosed for their spatio-temporal consistency. Key frames which are the characteristics frame of a particular video can be used to consolidate similar frames for two matching videos. Key frames are extracted by detecting shot boundaries measured using gray level changes in spatio-temporal part of the video and thereafter attaining their threshold. This is followed by application of video hashing technique to each and every frame. Matching of frames is achieved by local indexing method which is robust enough for different video transformations and also in the context of memory usage and computational period. The image based approach confronts with limitations such as vulnerability to attacks on video signals more likely in brightness/contrast, rotation, frame loss, addition of noise and spatio-temporal shift. [5]

C. Content- Based Video Copy Detection

Content-based video copy detection can be considered as an alternative approach to watermarking and Image based approach where their limitations such as ensuring legitimate content, illegal attacks and vulnerability to attacks on video signals are tackled. The underlying thesis of a content based copy detection system is that whenever a query video is encountered its fingerprint is generated foremost and then this is used to find a match for a fingerprint in a preformed video fingerprint database. If there is a similar fingerprint then the query video is a pirated version or illegal copy of a video otherwise not. Fig. 1 illustrates the step by step process involved. [5] The whole procedure can be briefed as:

1. Fingerprint Generation.
2. Fast search in video fingerprint database.
3. Decision making based on the search result.

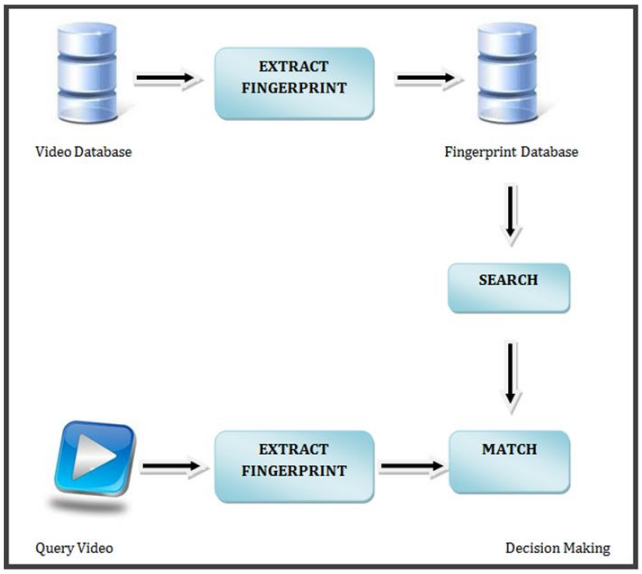

Figure 1: Content Based Video Copy Detection

Fingerprint Generation: A fingerprint generation algorithm should generate robust and discriminant fingerprints so that they can be made applicable in video applications. A variety of techniques has been adopted to extract these pertinent features and some of them are discussed in this section. Indyk et al has used temporal fingerprints based on shot boundaries of a video sequence. This technique has proved to be efficient for entire movie but is seen inefficient for short clips. A major advancement in this regard was made by Oostveen et al by making use of spatio-temporal fingerprint based on differential of luminance of partitioned grids in spatial and temporal regions. He presented the concept of hash function as a tool for video identification. B. Coskun et al [13] has proposed two robust hash algorithms which are based on Discrete Cosine transform [DCT] for video copy identification. These hash functions are more robust and random making it resistant to signal processing and transmission impairments and paving way in building database search, broadcast monitoring and watermarking applications of the video. It is realized that in spite of their robust nature DCT hash lacks security aspect as different video clips can have same hash value.

Hampapur and Bolle [6] have done significant work in this regard by comparing global descriptions of the video based on motion, color and spatio-temporal distribution of intensities. This ordinal measure proposal was actually put forward by Bhat and Nayar for computing image correspondences and then adapted by Mohan for video purposes. Studies have proved that ordinal measurements are robust to variety of resolutions, illumination shifts and

DOI: 03.LSCS.2013.5.47

display formats. Exact copy detection has been focused by Y. Li et al using compact binary signature involving color histograms. The use of local descriptors has proved better than ordinal measure which lacks robustness with respect to shifting or cropping of videos. Some signatures are also based on compact representation of the image content while limiting the correlation and redundancy between the features. M. Malekesmaeili et. al proposed an approach for generating spatio-temporal fingerprints denoted as TIRIs, Temporally Informative Representative Images [11]. Here performance was demonstrated by applying a simple image hashing technique on TIRIs of a video database.

Another fingerprinting algorithm which was proposed by Esmaeili et. al [12] to provide robustness in changing frame sizes of a video is TIRI-DCT algorithm. This was achieved using preprocessing step in this algorithm where a video is resized, divided into segments and changed to a standard format. The resulting images are called TIRI [Temporally Informative Representative Images] and are then applied with 2D-DCT [Discrete Cosine Transform] on overlapping blocks of size from each TIRI. The first horizontal and first vertical DCT coefficients are extracted from each block. Fig. 2 depicts the TIRI-DCT algorithm.

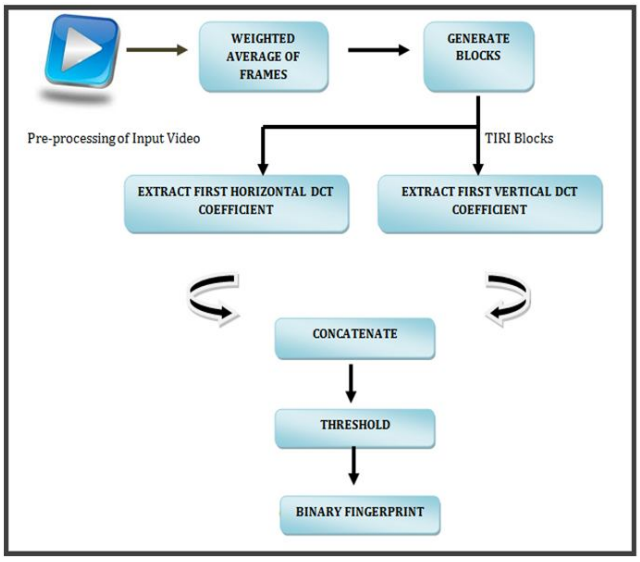

Figure 2: Schematic Representation of TIRI-DCT Algorithm

Different weight factors are incorporated such as constant, linear and exponential and these features are concatenated to form the feature vector which is later compared with a threshold to form the binary fingerprint. Another approach was proposed to concatenate TIRI-DCT with DWT to augment the efficiency of the system and gather more prominent features.

Searching in Video Database: After the generation of Fingerprints these are store in a database. Whenever a query video comes forth its fingerprint is first extracted and then searched in the fingerprint database for the closest fingerprint to the query video fingerprint. Fingerprints of two different copies of the same video will have similar fingerprints but not identical. This would help to find from which video this has been copied. Different approaches have been put forward in this regard and two of them which gained widespread acceptance were Inverted file based similarity search and cluster based approach.

In Inverted file based similarity search each fingerprint is divided into overlapping m bit blocks called words which are used to create inverted files from fingerprints [1]. These fingerprints are of equal length and can be represented as a table of size 2m×n where n is the no of words in a fingerprint of length L. Whenever there is a query fingerprint firstly it is divided into words and then compared to all fingerprints starting with the same word. The indices are found based on the corresponding entry in the first column of the inverted file table. The Euclidean distance is calculated and if it is less than some predefined threshold which is the median of spatio-temporal features, there is a match. If no such match is found, the procedure is repeated for the fingerprints that have exactly the same second word as the query's second word otherwise until up to the last word is checked for [2].

In Cluster based similarity search each fingerprint is assigned to one cluster and these fingerprints in the database will be clustered into groups. The cluster head closest to the query is determined and then all the fingerprints belonging to this cluster are searched to find a match i.e. which has the minimum Euclidean distance from the query. If a match is not found the cluster head second closest to query is determined and this repeats until a match is found. The cluster head must be chosen in such a way that even a minute change in fingerprint does not forward the fingerprint being assigned to different cluster [2].

Decision Making Based on Search Result: The decision making is performed on the basis of the search result. If there is a matching fingerprint in the video database then the query video is an illegal version otherwise it is not a copied video.

V. Comparison between Video Copy Detection Techniques

This section provides a comparison on different video copy detection techniques previously discussed in section IV. The comparisons were done based on the procedure used, attacks on different techniques and modification of content of the video. The following Table 1 illustrates their differences.

VI. Applications of Video Fingerprinting

Video fingerprinting technology identifies video content accurately and efficiently which paves way to many practical applications. Some of the applications include video Content Registration (video content described by use of metadata that binds with video fingerprint), Video Content Filtering (use of content filters to identify and filter out unauthorized copyrighted video content), Video Content tracking (where the video content is distributed and how many people have watched it), Broadcast Monitoring (can find out where and when a program has been broadcasted and how many times), Contextual advertising (can find out exactly what content is being consumed providing context for relevant advertisements), Video asset management (copies and edits of a video can easily be identified based on their unique

DOI: 03.LSCS.2013.5.47

identifier) and content-Based video search (find copies of video content, partial or whole, transformed or unaltered). [9]

TABLE 1. COMPARISON BETWEEN DIFFERENT VIDEO COPY DETECTION TECHNIQUES

Watermarking approach	Image based approach	Content based approach
A unique identifier embedded to the video imperceptible to human eye.	Based on key frame extractions which are characteristics frames of a video.	Unique features of video extracted and compared to reference fingerprints stored in a database.
Video content modified by addition of identification data.	Video content not modified.	Video content not affected
Precise identification of each piece of content allowed	Robust to significant video transformations.	Works out for legacy content
Standalone	Based on matching of individual key frames.	Connection to database required
Should embed information prior to release of video and is subjected to attacks such as watermarks compromised.	Is more subjected to attacks on video signals such as changes in brightness, rotation etc.	Capable of identifying a video which is under circulation and is more tolerable to video attacks.

VI. FUTURE WORK AND CONCLUSION

This paper imparts a brief idea on necessity of video copy detection techniques with a detailed description on some of the techniques presently adopted. The paper also emphasized on content based video copy detection describing the various steps involved. Also a comparison on different video copy detection techniques is also presented. One can conclude that the content based video copy detection techniques is more promising than the other techniques because of its ability to resist video attacks and also enable copy detection for videos even in circulation. On the contrary the watermarking approach and image based approach are vulnerable to attacks such as watermark compromise and attack in video signals respectively. So content based video copy detection is the best technique towards copy detection perspective and it is really worth researching because it still has the potential to be better. Also it has been realized that after generation of fingerprints the searching of the fingerprints from a large database consume considerable amount of time with the existing adopted techniques. So vital research should be done in this regard by incorporating more efficient mechanisms to deal with such large data and thus making the search faster.

ACKNOWLEDGMENT

We are greatly indebted to the college management and the faculty members for providing necessary facilities and hardware along with timely guidance and suggestions for implementing this work.

REFERENCES

[1] Diana Andrushia.A, Shilpa Harikumar, "Content Based Fingerprinting for VideoCopyright Detection System," International Journal of Emerging Technology and Advanced Engineering - ISSN 2250-2459, Volume 2, Issue 8, August 2012.

[2] S.Devi, N. Vishwanath, "A Robust Video Copy Detection Using TIRI-DCT Fingerprints and DWT Statistical Features," Asian Journal of Information Technology, ISSN 1682-3915, 2012.

[3] R. Gnana Rubini, P.Tamije Selvy, P. Anantha Prabha, "Performance Analysis of Fingerprinting Extraction Algorithm in Video Copy Detection System," International Journal of Computer Science, Engineering and Information Technology (IJCSEIT), Vol.2, No.1, February 2012

[4] Anindya Sarkar , Vishwakarma Singh , Pratim Ghosh , B. S. Manjunath , Ambuj Singh, "Efficient and Robust Detection of Duplicate Videos in a Large Database", IEEE Transactions on Circuits and systems for Video Technology, ISSN 1051-8215, June 2010

[5] M.Ramya , R.Kanthvel, "Efficient and Scalable Content- based Video Copy Detection System," International Conference on Recent Trends in Computational Methods, Communication and Controls (ICON3C 2012) Proceedings published in International Journal of Computer Applications (IJCA).

[6] Hampapur, A., Hyun, K.-H. and Bolle, R. M., "Comparison of sequence matching techniques for video copy detection," Proc. SPIE, Storage and Retrieval for Media Databases, vol. 4676, pp. 194-201, Jan. 2002.

[7] A. Joly, C. Frelicot and O. Buisson, "Feature statistical retrieval applied to content-based copy identification" International Conference on Image Processing, 2004.

[8] Abbass S. Abbass, Aliaa A. A. Youssif, Atef Z. Ghalwash "Hybrid-Based Compressed Domain Video Fingerprinting Technique" Computer and Information Science, Vol. 5, No. 5; 2012 ISSN 1913-8989 E-ISSN 1913-8997

[9] Jian Lu , Video fingerprinting for copy identification: from research to industry applications - Proceedings of SPIE - Media Forensics and Security XI, Vol. 7254, January 2009

[10] Wei-Lun Chao, "Comparison of Video Copy Detection techniques: The Robustness against Distortion and attacking," Tutorial Report, NTU, 2009

[11] M. Malekesmaeili, M. Fatourechi, and R. K. Ward, "Video copy detection using temporally informative representative images," in Proc. International Conference Machine Learning and Applications, Dec.2009, pp. 69–74.

[12] Mani Malek Esmaeili, Mehrdad Fatourechi, and Rabab Kreidieh Ward,Fellow, IEEE, "A robust and fast video copy detection system using content-based fingerprinting," IEEE Trans. On Information Forensics and Security, vol. 6, no. 1, Mar. 2011.

[13] B. Coskun, B. Sankur, and N. Memon, "Spatiotemporal transform based video hashing," IEEE Trans. Multimedia, vol. 8, no. 6, pp.1190–1208, Dec. 2006.

[14] C. De Roover, C. De Vleeschouwer, F. Lefebvre, and B. Macq, "Robust video hashing based on radial projections of key frames," IEEE Trans. Signal Process, vol. 53, no. 10, pp. 4020–4037, Oct. 2005.

[15] Wei- Lun Chao, "Introduction to video Fingerprinting," Tutorial Report, DISP lab, NTU, 2009.

DOI: 03.LSCS.2013.5. 47